Full wwPDB NMR Structure Validation Report (i)

May 28, 2020 – 11:58 pm BST

PDB ID	:	2MKH
Title	:	Solution structure of tandem RRM domains of cytoplasmic polyadenylation element binding protein 1 (CPEB1) in free state
Authors	:	Afroz, T.; Skrisovska, L.; Belloc, E.; Boixet, J.G.; Mendez, R.; Allain, F.H.-T.
Deposited on	:	2014-02-07

This is a Full wwPDB NMR Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

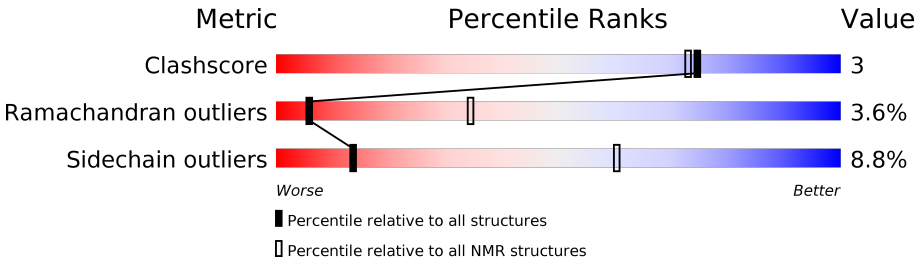
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 66%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	217	87% 9%

2 Ensemble composition and analysis

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:218-A:304, A:309-A:429 (208)	1.04	1

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 4 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 4, 5, 7, 8, 10, 14, 16
2	3, 9, 13, 15, 19
3	6, 11, 20
4	17, 18
Single-model clusters	12

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3360 atoms, of which 1682 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cytoplasmic polyadenylation element-binding protein 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	212	3360	1082	1682	289	299	8	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	218	MET	-	EXPRESSION TAG	UNP Q9BZB8

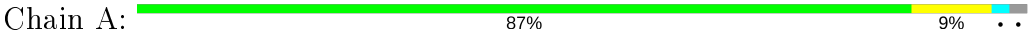
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:



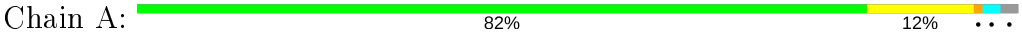
4.2 Scores per residue for each member of the ensemble

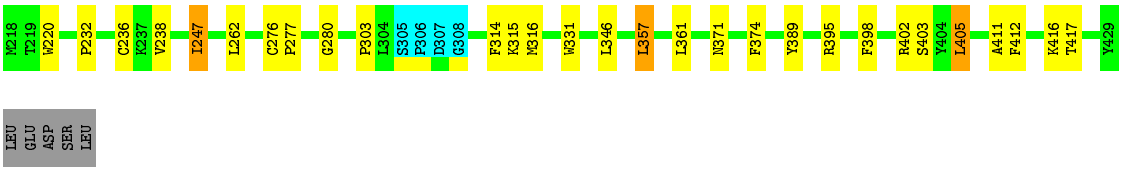
Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:

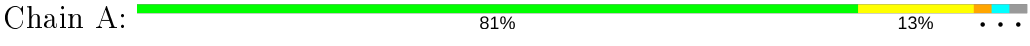


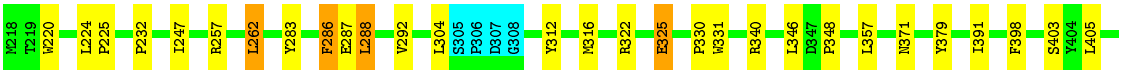


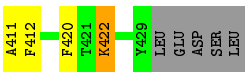
4.2.2 Score per residue for model 2

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:



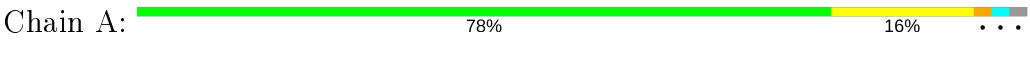


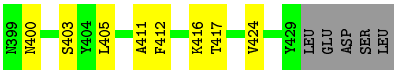


4.2.3 Score per residue for model 3

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:

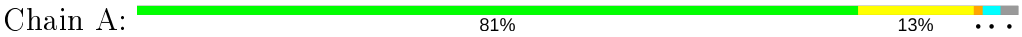


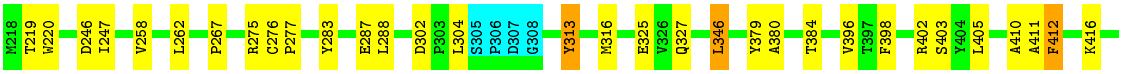


4.2.4 Score per residue for model 4

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:

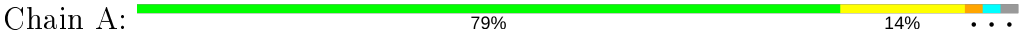


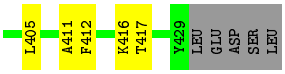


4.2.5 Score per residue for model 5

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:

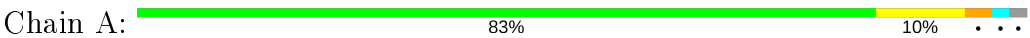


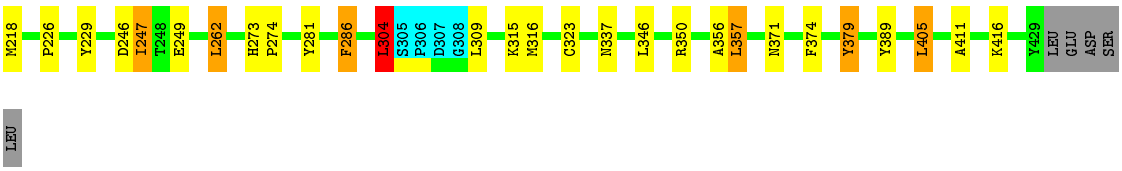


4.2.6 Score per residue for model 6

• Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:

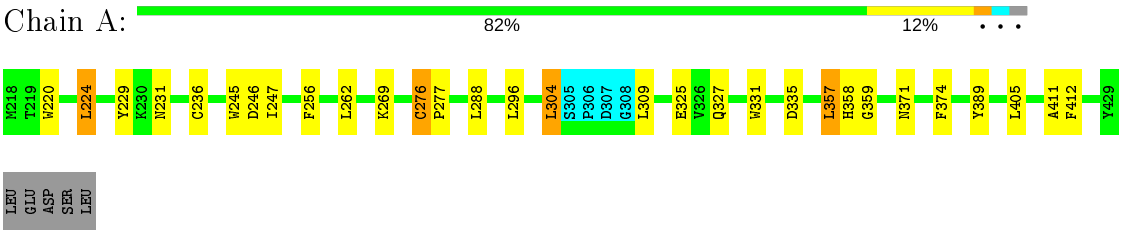




4.2.7 Score per residue for model 7

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

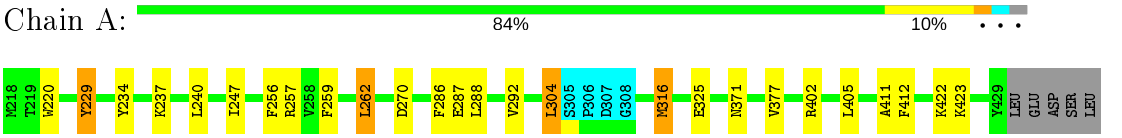
Chain A:



4.2.8 Score per residue for model 8

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

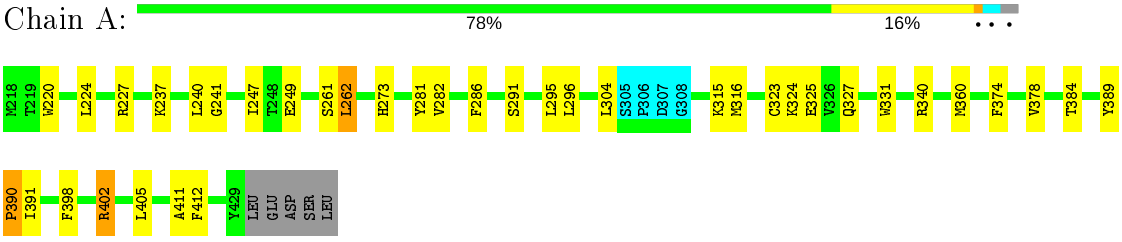
Chain A:



4.2.9 Score per residue for model 9

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

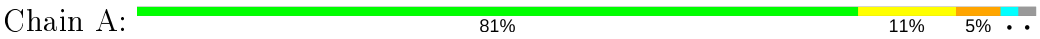
Chain A:

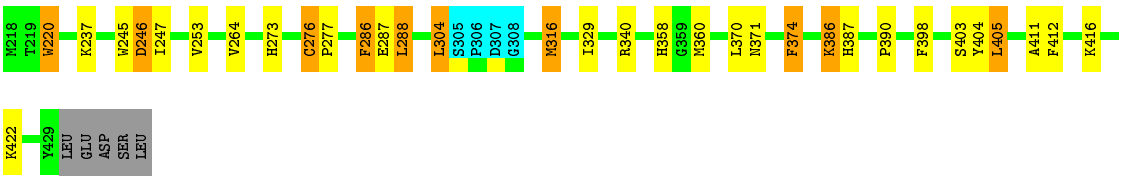


4.2.10 Score per residue for model 10

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:

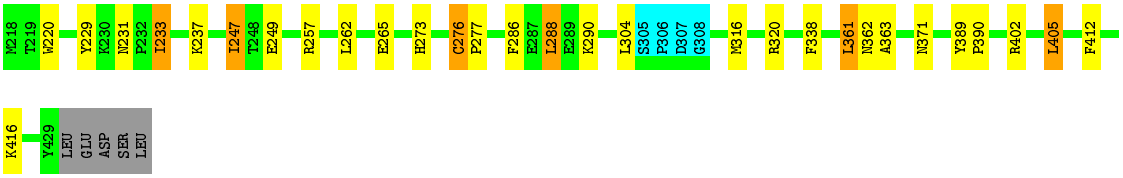




4.2.11 Score per residue for model 11

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

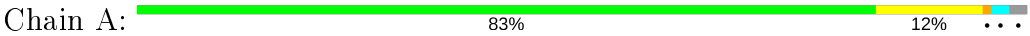
Chain A:

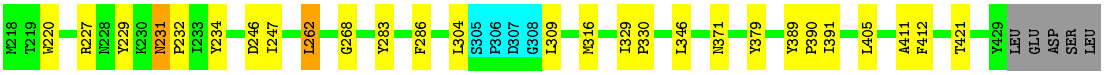


4.2.12 Score per residue for model 12

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:

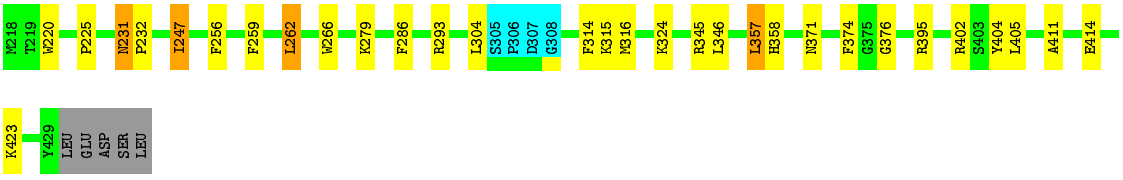




4.2.13 Score per residue for model 13

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

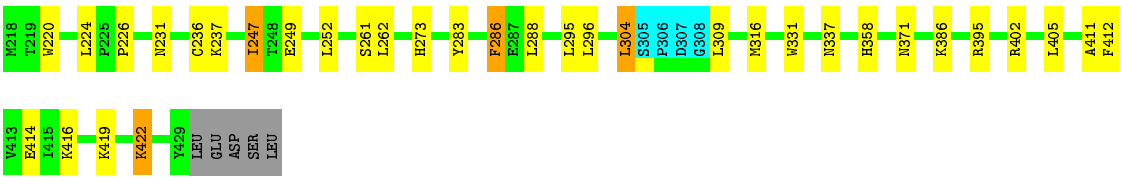
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

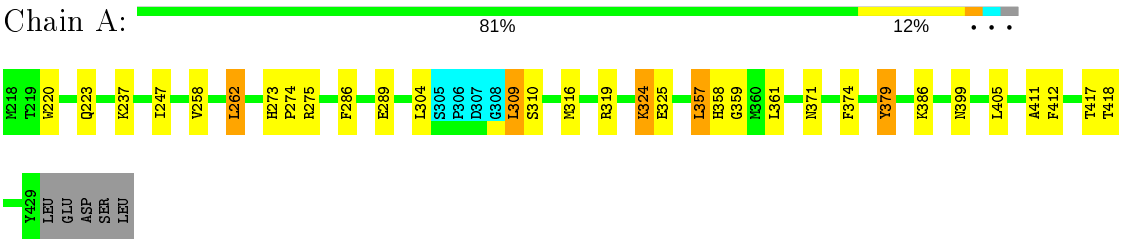
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

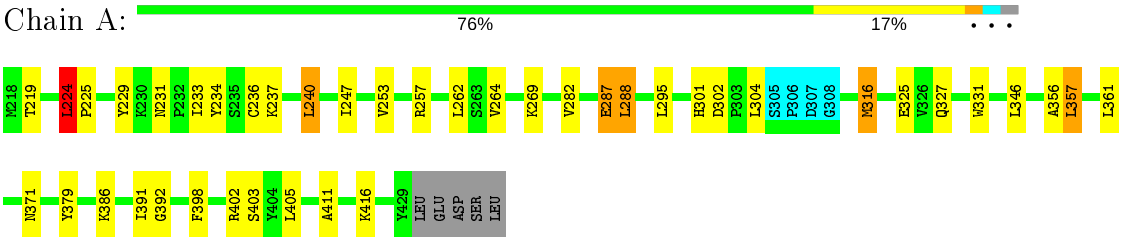
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

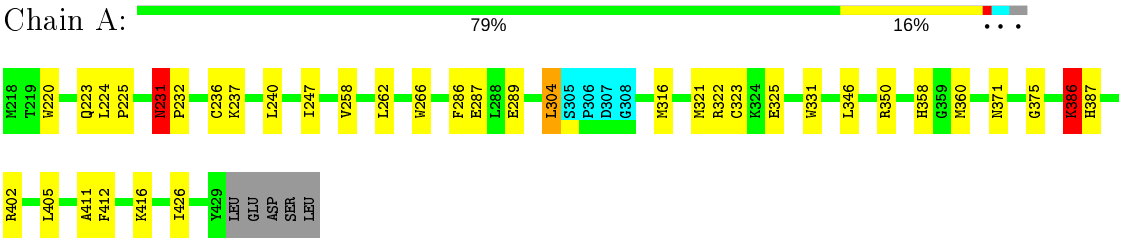
Chain A:



4.2.17 Score per residue for model 17

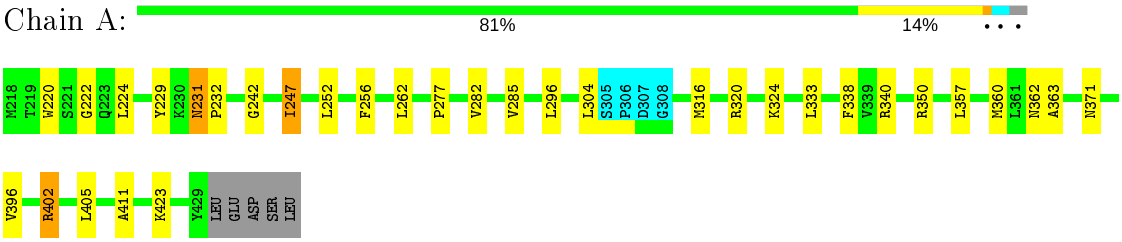
- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1

Chain A:



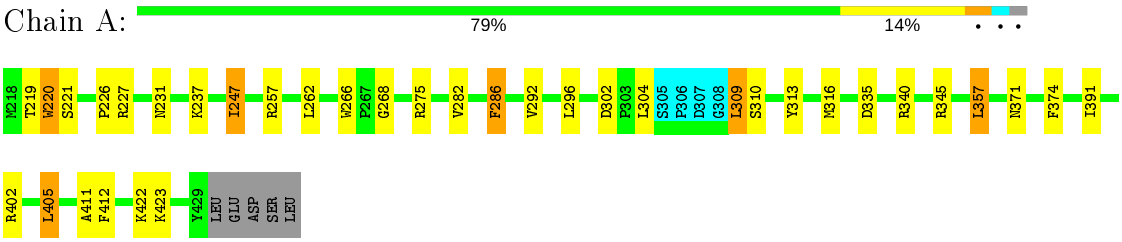
4.2.18 Score per residue for model 18

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1



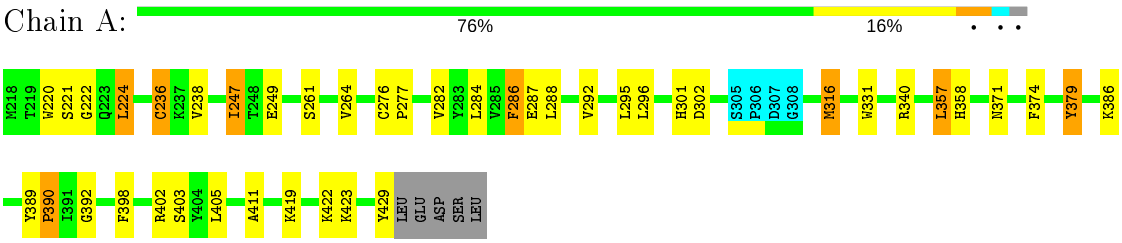
4.2.19 Score per residue for model 19

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1



4.2.20 Score per residue for model 20

- Molecule 1: Cytoplasmic polyadenylation element-binding protein 1



5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing.*

Of the 30 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy.*

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
AMBER	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	2037
Number of shifts mapped to atoms	2037
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	66%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1653	1663	1660	11±4
All	All	33060	33260	33200	213

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:357:LEU:H	1:A:357:LEU:HD13	0.78	1.39	1	2
1:A:346:LEU:H	1:A:346:LEU:HD22	0.64	1.52	16	1
1:A:288:LEU:HD23	1:A:288:LEU:H	0.62	1.54	4	3
1:A:247:ILE:H	1:A:247:ILE:CD1	0.62	2.07	19	6
1:A:247:ILE:H	1:A:247:ILE:HD13	0.62	1.54	13	2
1:A:247:ILE:HD13	1:A:247:ILE:H	0.61	1.54	19	3
1:A:357:LEU:HD13	1:A:357:LEU:H	0.59	1.57	19	4
1:A:276:CYS:H	1:A:277:PRO:HD2	0.59	1.58	7	2
1:A:253:VAL:CG2	1:A:264:VAL:HG21	0.59	2.28	16	2
1:A:276:CYS:H	1:A:277:PRO:CD	0.58	2.10	7	1
1:A:261:SER:H	1:A:295:LEU:HD11	0.57	1.58	14	2
1:A:252:LEU:H	1:A:252:LEU:HD23	0.57	1.58	18	2
1:A:231:ASN:N	1:A:232:PRO:CD	0.55	2.70	12	3
1:A:304:LEU:C	1:A:304:LEU:HD22	0.54	2.23	14	1
1:A:398:PHE:CE1	1:A:403:SER:HB3	0.54	2.38	20	2
1:A:276:CYS:SG	1:A:277:PRO:HD2	0.54	2.43	10	2
1:A:233:ILE:HG22	1:A:234:TYR:H	0.54	1.61	16	1
1:A:357:LEU:CD1	1:A:357:LEU:H	0.54	2.13	1	3
1:A:288:LEU:H	1:A:288:LEU:HD23	0.54	1.63	8	3
1:A:286:PHE:H	1:A:288:LEU:CD2	0.53	2.17	14	1
1:A:262:LEU:HD13	1:A:262:LEU:H	0.53	1.63	13	2
1:A:357:LEU:H	1:A:357:LEU:HD23	0.53	1.63	2	1
1:A:233:ILE:H	1:A:233:ILE:HD13	0.53	1.62	11	1
1:A:224:LEU:N	1:A:225:PRO:CD	0.53	2.71	3	5
1:A:398:PHE:CD2	1:A:403:SER:CB	0.52	2.93	16	2
1:A:356:ALA:C	1:A:357:LEU:HD13	0.52	2.24	6	2
1:A:325:GLU:CD	1:A:325:GLU:H	0.52	2.09	15	1
1:A:262:LEU:N	1:A:262:LEU:HD13	0.51	2.20	8	2
1:A:247:ILE:N	1:A:247:ILE:HD13	0.51	2.19	19	1
1:A:253:VAL:HG22	1:A:264:VAL:HG21	0.51	1.81	16	1
1:A:236:CYS:HB2	1:A:331:TRP:CG	0.51	2.41	17	4
1:A:286:PHE:CE1	1:A:292:VAL:HG22	0.51	2.40	19	2
1:A:316:MET:SD	1:A:316:MET:N	0.51	2.84	20	2
1:A:261:SER:H	1:A:295:LEU:CD1	0.51	2.19	14	1
1:A:276:CYS:N	1:A:277:PRO:CD	0.51	2.73	1	3
1:A:379:TYR:CG	1:A:380:ALA:N	0.51	2.79	3	2
1:A:262:LEU:HD22	1:A:262:LEU:C	0.50	2.26	12	1
1:A:288:LEU:CD2	1:A:288:LEU:H	0.50	2.18	4	1
1:A:236:CYS:HB3	1:A:331:TRP:CG	0.50	2.41	1	1
1:A:346:LEU:HD22	1:A:346:LEU:N	0.50	2.22	1	3
1:A:224:LEU:N	1:A:224:LEU:HD13	0.49	2.21	3	1
1:A:224:LEU:H	1:A:224:LEU:HD13	0.49	1.66	3	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:422:LYS:HE3	1:A:422:LYS:HA	0.49	1.85	14	1
1:A:345:ARG:HG3	1:A:346:LEU:H	0.49	1.68	13	1
1:A:257:ARG:HA	1:A:262:LEU:HD11	0.48	1.85	2	2
1:A:286:PHE:CZ	1:A:292:VAL:HG22	0.48	2.43	2	1
1:A:379:TYR:CD1	1:A:379:TYR:C	0.48	2.86	15	4
1:A:357:LEU:HD12	1:A:357:LEU:O	0.48	2.09	15	1
1:A:224:LEU:N	1:A:225:PRO:HD3	0.48	2.23	3	1
1:A:267:PRO:HD3	1:A:283:TYR:CE2	0.48	2.44	4	1
1:A:357:LEU:HD13	1:A:357:LEU:N	0.48	2.24	16	2
1:A:304:LEU:N	1:A:304:LEU:HD12	0.48	2.24	18	2
1:A:262:LEU:C	1:A:262:LEU:HD22	0.48	2.29	15	2
1:A:233:ILE:HG22	1:A:234:TYR:N	0.48	2.23	16	1
1:A:220:TRP:CZ3	1:A:222:GLY:CA	0.47	2.96	3	1
1:A:333:LEU:HD23	1:A:333:LEU:O	0.47	2.08	18	1
1:A:422:LYS:HG3	1:A:423:LYS:H	0.47	1.70	8	3
1:A:309:LEU:HD13	1:A:310:SER:H	0.47	1.69	15	1
1:A:405:LEU:H	1:A:405:LEU:HD13	0.47	1.70	19	1
1:A:321:MET:C	1:A:323:CYS:H	0.47	2.12	17	1
1:A:361:LEU:HD13	1:A:362:ASN:N	0.47	2.25	11	2
1:A:402:ARG:HD2	1:A:402:ARG:H	0.47	1.70	1	1
1:A:398:PHE:CD1	1:A:403:SER:HB2	0.47	2.45	3	1
1:A:357:LEU:HD22	1:A:359:GLY:H	0.47	1.68	7	1
1:A:370:LEU:HD13	1:A:374:PHE:CE2	0.47	2.45	10	1
1:A:338:PHE:HB2	1:A:363:ALA:HB1	0.46	1.87	11	2
1:A:374:PHE:CZ	1:A:376:GLY:HA3	0.46	2.46	13	1
1:A:224:LEU:H	1:A:224:LEU:HD22	0.46	1.69	20	1
1:A:249:GLU:H	1:A:249:GLU:CD	0.46	2.13	20	1
1:A:309:LEU:H	1:A:309:LEU:HD23	0.46	1.70	14	1
1:A:224:LEU:H	1:A:224:LEU:CD2	0.46	2.23	20	1
1:A:247:ILE:N	1:A:247:ILE:CD1	0.46	2.78	19	1
1:A:231:ASN:H	1:A:232:PRO:CD	0.46	2.24	17	1
1:A:220:TRP:CH2	1:A:222:GLY:CA	0.45	2.99	3	1
1:A:220:TRP:CZ3	1:A:222:GLY:N	0.45	2.84	3	1
1:A:405:LEU:HD22	1:A:405:LEU:C	0.45	2.32	6	3
1:A:331:TRP:CZ3	1:A:391:ILE:HB	0.45	2.47	2	1
1:A:398:PHE:CD1	1:A:403:SER:CB	0.45	3.00	4	3
1:A:286:PHE:O	1:A:287:GLU:C	0.45	2.55	20	1
1:A:426:ILE:HD12	1:A:426:ILE:N	0.45	2.27	17	1
1:A:346:LEU:CD2	1:A:346:LEU:N	0.45	2.80	12	1
1:A:346:LEU:HD11	1:A:379:TYR:CE1	0.44	2.47	4	1
1:A:402:ARG:H	1:A:402:ARG:HD2	0.44	1.72	9	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:309:LEU:HD12	1:A:310:SER:N	0.44	2.26	19	1
1:A:346:LEU:CD1	1:A:379:TYR:CD1	0.44	3.01	2	1
1:A:389:TYR:N	1:A:390:PRO:CD	0.44	2.81	9	1
1:A:346:LEU:HD11	1:A:379:TYR:CD1	0.44	2.48	16	1
1:A:316:MET:SD	1:A:316:MET:C	0.44	2.96	16	2
1:A:324:LYS:HA	1:A:324:LYS:HE2	0.44	1.90	15	1
1:A:346:LEU:H	1:A:346:LEU:CD2	0.44	2.25	16	1
1:A:241:GLY:HA3	1:A:281:TYR:CB	0.43	2.43	9	1
1:A:398:PHE:CE1	1:A:403:SER:HB2	0.43	2.48	10	1
1:A:325:GLU:H	1:A:325:GLU:CD	0.43	2.17	2	1
1:A:414:GLU:CD	1:A:414:GLU:N	0.43	2.71	14	1
1:A:224:LEU:H	1:A:225:PRO:CD	0.43	2.25	16	1
1:A:313:TYR:CE2	1:A:327:GLN:HB3	0.43	2.48	4	1
1:A:238:VAL:HG11	1:A:286:PHE:CE1	0.43	2.47	20	1
1:A:386:LYS:HA	1:A:386:LYS:HE3	0.43	1.89	5	1
1:A:224:LEU:O	1:A:226:PRO:HD3	0.43	2.14	14	1
1:A:304:LEU:N	1:A:304:LEU:CD1	0.43	2.82	19	1
1:A:240:LEU:C	1:A:240:LEU:HD13	0.43	2.34	16	2
1:A:414:GLU:HG2	1:A:423:LYS:HE2	0.43	1.91	13	1
1:A:361:LEU:C	1:A:361:LEU:HD13	0.43	2.35	16	1
1:A:301:HIS:CG	1:A:302:ASP:N	0.43	2.86	16	2
1:A:262:LEU:CD2	1:A:262:LEU:N	0.43	2.81	18	1
1:A:247:ILE:CD1	1:A:247:ILE:H	0.42	2.26	1	1
1:A:302:ASP:C	1:A:304:LEU:H	0.42	2.17	19	1
1:A:236:CYS:HB3	1:A:331:TRP:CB	0.42	2.44	20	1
1:A:247:ILE:HD13	1:A:247:ILE:N	0.42	2.29	1	1
1:A:405:LEU:C	1:A:405:LEU:HD22	0.42	2.34	1	1
1:A:374:PHE:CZ	1:A:402:ARG:HB2	0.42	2.48	13	1
1:A:288:LEU:N	1:A:288:LEU:HD23	0.42	2.25	4	1
1:A:240:LEU:HD13	1:A:240:LEU:C	0.42	2.35	8	1
1:A:291:SER:O	1:A:295:LEU:HG	0.42	2.14	9	1
1:A:286:PHE:CD1	1:A:286:PHE:N	0.42	2.87	10	1
1:A:245:TRP:CE3	1:A:321:MET:CE	0.42	3.03	3	1
1:A:234:TYR:CE1	1:A:330:PRO:HG2	0.42	2.50	12	1
1:A:314:PHE:CG	1:A:315:LYS:N	0.42	2.88	1	1
1:A:357:LEU:HB3	1:A:424:VAL:HG13	0.42	1.91	3	1
1:A:273:HIS:N	1:A:274:PRO:HD3	0.42	2.30	6	1
1:A:262:LEU:CB	1:A:286:PHE:HB3	0.42	2.45	6	1
1:A:234:TYR:CD2	1:A:292:VAL:CG1	0.42	3.03	8	1
1:A:416:LYS:HE2	1:A:416:LYS:HA	0.42	1.91	11	1
1:A:223:GLN:HA	1:A:417:THR:HG22	0.42	1.92	5	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:331:TRP:CH2	1:A:391:ILE:HB	0.42	2.50	9	1
1:A:314:PHE:CD1	1:A:315:LYS:N	0.41	2.85	1	3
1:A:386:LYS:HE2	1:A:387:HIS:CE1	0.41	2.50	10	2
1:A:360:MET:O	1:A:362:ASN:N	0.41	2.53	18	1
1:A:309:LEU:H	1:A:309:LEU:CD2	0.41	2.28	14	1
1:A:240:LEU:C	1:A:240:LEU:CD1	0.41	2.89	16	1
1:A:420:PHE:CZ	1:A:422:LYS:HD3	0.41	2.50	2	1
1:A:374:PHE:CG	1:A:375:GLY:N	0.41	2.88	5	1
1:A:257:ARG:O	1:A:257:ARG:CG	0.41	2.68	19	1
1:A:304:LEU:HD22	1:A:304:LEU:C	0.41	2.35	7	1
1:A:262:LEU:HA	1:A:286:PHE:CB	0.41	2.45	15	1
1:A:329:ILE:O	1:A:329:ILE:HG23	0.41	2.16	10	1
1:A:389:TYR:N	1:A:390:PRO:HD2	0.41	2.31	20	1
1:A:262:LEU:HD13	1:A:262:LEU:N	0.41	2.30	2	1
1:A:422:LYS:CE	1:A:422:LYS:HA	0.41	2.46	14	1
1:A:219:THR:C	1:A:221:SER:H	0.41	2.19	19	1
1:A:410:ALA:C	1:A:412:PHE:H	0.41	2.19	4	1
1:A:379:TYR:C	1:A:379:TYR:CD1	0.41	2.94	16	1
1:A:357:LEU:H	1:A:357:LEU:CD1	0.40	2.27	13	1
1:A:417:THR:HG22	1:A:418:THR:H	0.40	1.77	15	1
1:A:237:LYS:HE2	1:A:283:TYR:CD1	0.40	2.52	14	1
1:A:358:HIS:CE1	1:A:360:MET:HG3	0.40	2.52	17	1
1:A:245:TRP:O	1:A:246:ASP:C	0.40	2.59	10	1

5.2 Torsion angles (i)

5.2.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	206/217 (95%)	168±4 (82±2%)	30±3 (15±1%)	7±3 (4±1%)	6	34
All	All	4120/4340 (95%)	3368 (82%)	605 (15%)	147 (4%)	6	34

All 55 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	411	ALA	19
1	A	224	LEU	7
1	A	390	PRO	6
1	A	246	ASP	6
1	A	287	GLU	6
1	A	386	LYS	6
1	A	288	LEU	5
1	A	391	ILE	4
1	A	258	VAL	4
1	A	273	HIS	4
1	A	325	GLU	4
1	A	231	ASN	4
1	A	282	VAL	4
1	A	219	THR	3
1	A	276	CYS	3
1	A	286	PHE	3
1	A	268	GLY	3
1	A	247	ILE	3
1	A	232	PRO	3
1	A	323	CYS	2
1	A	358	HIS	2
1	A	237	LYS	2
1	A	374	PHE	2
1	A	392	GLY	2
1	A	346	LEU	2
1	A	360	MET	2
1	A	322	ARG	2
1	A	375	GLY	2
1	A	303	PRO	2
1	A	289	GLU	2
1	A	226	PRO	2
1	A	269	LYS	2
1	A	222	GLY	2
1	A	348	PRO	1
1	A	242	GLY	1
1	A	324	LYS	1
1	A	344	GLN	1
1	A	277	PRO	1
1	A	357	LEU	1
1	A	227	ARG	1
1	A	280	GLY	1
1	A	220	TRP	1
1	A	309	LEU	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	275	ARG	1
1	A	361	LEU	1
1	A	378	VAL	1
1	A	270	ASP	1
1	A	223	GLN	1
1	A	225	PRO	1
1	A	304	LEU	1
1	A	320	ARG	1
1	A	221	SER	1
1	A	274	PRO	1
1	A	350	ARG	1
1	A	359	GLY	1

5.2.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	181/189 (96%)	165±3 (91±1%)	16±3 (9±1%)	13 60
All	All	3620/3780 (96%)	3300 (91%)	320 (9%)	13 60

All 77 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	405	LEU	20
1	A	316	MET	19
1	A	247	ILE	18
1	A	262	LEU	17
1	A	220	TRP	17
1	A	371	ASN	17
1	A	412	PHE	15
1	A	304	LEU	15
1	A	402	ARG	11
1	A	357	LEU	10
1	A	286	PHE	9
1	A	416	LYS	9
1	A	237	LYS	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	229	TYR	6
1	A	231	ASN	6
1	A	296	LEU	6
1	A	327	GLN	5
1	A	379	TYR	5
1	A	358	HIS	4
1	A	325	GLU	4
1	A	309	LEU	4
1	A	340	ARG	4
1	A	361	LEU	4
1	A	249	GLU	4
1	A	324	LYS	3
1	A	422	LYS	3
1	A	288	LEU	3
1	A	224	LEU	3
1	A	384	THR	3
1	A	396	VAL	3
1	A	266	TRP	3
1	A	295	LEU	3
1	A	238	VAL	2
1	A	374	PHE	2
1	A	417	THR	2
1	A	335	ASP	2
1	A	337	ASN	2
1	A	389	TYR	2
1	A	256	PHE	2
1	A	290	LYS	2
1	A	240	LEU	2
1	A	315	LYS	2
1	A	419	LYS	2
1	A	236	CYS	2
1	A	259	PHE	2
1	A	386	LYS	2
1	A	261	SER	1
1	A	230	LYS	1
1	A	400	ASN	1
1	A	282	VAL	1
1	A	330	PRO	1
1	A	399	ASN	1
1	A	423	LYS	1
1	A	273	HIS	1
1	A	329	ILE	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	319	ARG	1
1	A	285	VAL	1
1	A	227	ARG	1
1	A	279	LYS	1
1	A	245	TRP	1
1	A	378	VAL	1
1	A	264	VAL	1
1	A	281	TYR	1
1	A	223	GLN	1
1	A	257	ARG	1
1	A	284	LEU	1
1	A	345	ARG	1
1	A	302	ASP	1
1	A	395	ARG	1
1	A	246	ASP	1
1	A	233	ILE	1
1	A	218	MET	1
1	A	265	GLU	1
1	A	287	GLU	1
1	A	292	VAL	1
1	A	377	VAL	1
1	A	421	THR	1

5.2.3 RNA

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.4 Carbohydrates

There are no carbohydrates in this entry.

5.5 Ligand geometry

There are no ligands in this entry.

5.6 Other polymers

There are no such molecules in this entry.

5.7 Polymer linkage issues

There are no chain breaks in this entry.

6 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 66% for the well-defined parts and 66% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *CPEB1RM12free*

6.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	2037
Number of shifts mapped to atoms	2037
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	11

6.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	152	2.98 ± 0.14	Should be applied
$^{13}C_{\beta}$	167	3.06 ± 0.13	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	195	-0.12 ± 0.17	None needed (< 0.5 ppm)

6.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 66%, i.e. 1715 atoms were assigned a chemical shift out of a possible 2608. 31 out of 38 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	670/1012 (66%)	339/402 (84%)	144/416 (35%)	187/194 (96%)
Sidechain	863/1341 (64%)	529/794 (67%)	321/482 (67%)	13/65 (20%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	182/255 (71%)	112/135 (83%)	66/112 (59%)	4/8 (50%)
Overall	1715/2608 (66%)	980/1331 (74%)	531/1010 (53%)	204/267 (76%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 66%, i.e. 1739 atoms were assigned a chemical shift out of a possible 2642. 31 out of 38 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	682/1030 (66%)	345/409 (84%)	147/424 (35%)	190/197 (96%)
Sidechain	875/1357 (64%)	537/804 (67%)	325/488 (67%)	13/65 (20%)
Aromatic	182/255 (71%)	112/135 (83%)	66/112 (59%)	4/8 (50%)
Overall	1739/2642 (66%)	994/1348 (74%)	538/1024 (53%)	207/270 (77%)

6.1.4 Statistically unusual chemical shifts (i)

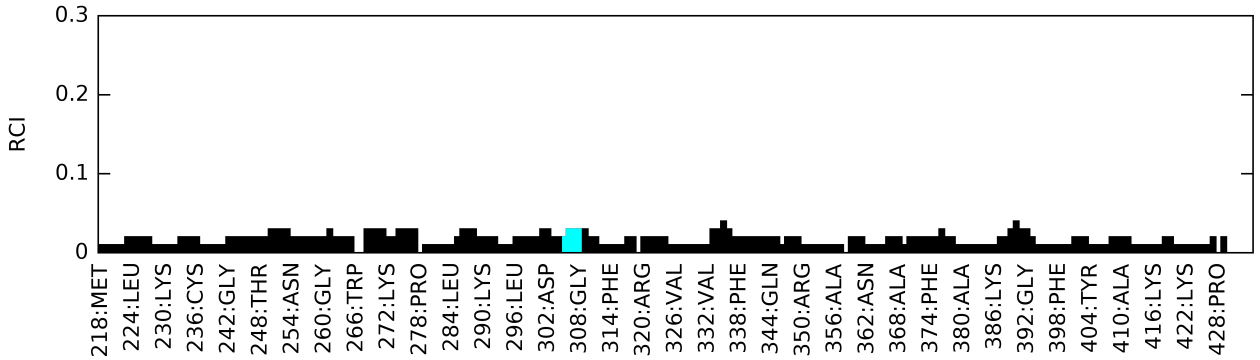
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	369	ILE	CG1	12.05	36.54 – 18.94	-8.9
1	A	237	LYS	HG2	3.61	2.67 – 0.07	8.6
1	A	237	LYS	HG3	3.61	2.76 – -0.04	8.0
1	A	390	PRO	HG3	-0.54	3.56 – 0.26	-7.4
1	A	390	PRO	HD3	0.81	5.52 – 1.72	-7.4
1	A	390	PRO	HD2	1.01	5.45 – 1.85	-7.3
1	A	311	GLU	CG	27.27	42.24 – 29.94	-7.2
1	A	331	TRP	HB2	0.91	4.94 – 1.44	-6.5
1	A	390	PRO	HG2	-0.05	3.48 – 0.38	-6.4
1	A	331	TRP	HB3	0.91	4.94 – 1.34	-6.2
1	A	373	LEU	HB2	-0.10	3.32 – -0.08	-5.1

6.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:


RCI
0.3
0.2
0.1
0
218:MET
224:LEU
230:LYS
236:CYS
242:GLY
248:THR
254:ASN
260:GLY
266:TRP
272:LYS
278:PRO
284:LEU
290:LYS
296:LEU
302:ASP
308:GLY
314:PHE
320:ARG
326:VAL
332:VAL
338:PHE
344:GLN
350:ARG
356:ALA
362:ASN
368:ALA
374:PHE
380:ALA
386:LYS
392:GLY
398:PHE
404:TYR
410:ALA
416:LYS
422:LYS
428:PRO